Chapter 3: Profit, Output Decisions, and Market Supply

I. Introduction: Production for Sale as a Utility Maximizing Activity

The previous chapter is a bit longer than most other chapters in this text because it laid the foundations for much of the rest of the book. It articulates and models the utility maximizing model of volition. It also developed a set of geometric and mathematical approaches to modelling that can be applied to a variety of choice settings as the remainder of the book will illustrate. Chapter 2 demonstrated how they could be applied to develop a series of increasingly general models of consumer behavior, from relatively simple one-choice at a time models to ones where many choices are made simultaneously. The rest of the book is grounded in the ideas and mathematical approaches used in Chapter 2 to develop the neoclassical theory of the demand for final goods and services (as opposed to intermediate goods and inputs).

This chapter focuses on how similar ideas and mathematical approaches can be used to model the quantity of goods and services available for sale in markets. Chapter 4 will tackle production itself and markets for intermediate goods and inputs. Other extensions and applications are worked out in Parts II and III of the textbook. Part III uses the general approach of Chapter 2 to model social phenomena such as crime and politics that affect market outcomes, but are less central to—indeed largely absent from—most overviews of neoclassical price theory.

Utility Maximization and Supply

The utility maximizing models of the previous chapter are relevant for the analysis of production as well as consumption. It is the core model of volition used in microeconomics and other rational-choice based models in game theory, public choice, and sociological and anthropological theories grounded in methodological individualism. In this subsection, we begin by showing how the same framework can be used to ground the theory of supply.

Although the aim of both input supply and the production of goods for sale is usually assumed to be the income generated by those activities, from the perspective of the utility-maximizing model of volition, such income is pursued only because it tends to increase utility. Income tends to expand one's opportunity set, and a larger opportunity set tends to increase realized utility. It is for this reason that producers of products for sale in markets can be modeled as income or profit maximizers. It is a simplification justified by the fact that their ultimate end is still the maximization of utility and increases in income tend to increase utility. This simplification is most appropriate when obtaining income is the only way in which an activity increases utility.

This characterization of firm owners and persons looking for jobs is plausible in part because of specialization. In many cases, the income earned by owning a firm or selling one's services to firms is the primary way that one obtains both the necessities of life and other goods and services that make life more pleasant.

When markets first emerged it is likely that intermediate cases were commonplace in which persons produce some things for themselves (as with vegetable gardens or hunting) and sell or trade part of their produce to others for things that they do not produce. For example, gardeners may trade part of their garden's output to a neighboring blacksmith in exchange for metal shovels, hoes, gate-latches, or hinges that are more durable or useful than their wooden substitutes. The blacksmith, in turn, may produce nearly all of his or her metal products for sale, while retaining a minor subset for his or her own use. It is through such processes that specialization and markets are likely to have first emerged.

In well-developed commercial systems, production by economic organizations is entirely for sale to others and most people in a community, region, or country "hire themselves out for wages" that they use to purchase goods and services produced by others or other organizations. Others form (or inherit) economic organizations (firms) through which they garner income for purchasing goods and services.

Such firm owners may realize additional benefits from their commercial activities. They may, for example, produce or sell products that they personally enjoy, as owners of a bicycle or coffee shop often do. These and other direct benefits from ownership are largely ignored in neoclassical models, because realizing net-income is nearly always the primary reason that such organizations are founded and operated.

Such specialization (in labor supply or firm ownership) is facilitated by the dense and ubiquitous commercial networks that emerged first in the nineteenth century in what is currently referred to as "the West." As commerce expanded, farming diminished in economic importance, and jobs in industry and services became increasingly commonplace and eventually the main source of income for most people. It is the market relationships within such commercial societies that neoclassical economics emerged to explain.

The first models of supply—and the ones that have attracted most of the attention of textbook authors—all assume that organizations called "firms" engage in the production of goods for sale, and that firm owners do so only because of the net-income that they realize from selling the goods and services produced. Because more income is better than less income (other things being equal), these models all assume that the aim of production is to maximize the profits that firm owners realize from selling the output(s) produced.

Other things being equal, larger incomes, of course, are associated with higher levels of utility for both input providers and firm owners. Other sources of utility from ownership and employment are assumed to be negligible and so can be ignored for purposes of modeling both firm ownership and labor supply—although there are clearly cases where job satisfaction is a significant source of utility for some owners and suppliers of labor.

II. The Geometry of Supply Decisions by "Price Taking" Firms

Specialization and the Importance of Income

Individual and small groups that devoted most of their efforts to producing goods for sale tended to specialize, because specialization tended to reduce production costs. The specific skills (human capital) for particular forms of production were improved through practice (as the saying goes practice makes perfect) and the use of specialized capital goods (hammers, spinning wheels, looms, trucks, and so forth) further reduced their production costs. These cost savings, in turn, made it possible to sell their goods and services to others at a price that was below the opportunity cost for non-specialists to produce the same products and provided net income (income above their opportunity cost in non-specialized production) that could be used to acquire goods and services from other specialized producers.

That such specialization occurred is evident in most languages, as terms for specialized activities emerged to describe them because they were so common: farmer, hunter, carpenter, spinner, clothier, shipper, and so forth.

In none of the professions listed would a firm or its employees be able to survive without selling their goods or services and using the proceeds to purchase foodstuffs from farms or markets for produce. Net income had to be realized to make their businesses or input supply feasible. And, other specialists must produce the goods and services required for a given specialist to feed, clothe, and house themselves and their families. Village craftsmen and craftswomen relied upon market transactions to meet their daily needs. They sold goods and services to farmers and others, and the proceeds were largely used to purchase other goods from other vendors.

In the days when village and town markets first emerged, producers and sellers were often single persons or small organizations that were largely managed and staffed by family members. As somewhat larger organizations emerged, they often remained family-based proprietorships, although with more non-family members employed by their firms. Until the mid-nineteenth century, such firms were often similar in size and similarly limited in their ability to increase their scale of operation. Thus, when demand expanded, additional supply would be provided by additional firms (e.g. "entry").

This is the historical basis of the Marshallian (1890) theory of long run supply, a topic taken up toward the end of this chapter. Most economic enterprises in the late nineteenth and early twentieth century were relatively small organizations in which it was relatively easy to monitor employees and assure that they remained "on task" during the working hours—which were often quite long during the 19th century. As management methods improved, larger firms could be (almost) equally well managed and so firm size tended to expand during the twentieth century. However, it is Marshall's (1890) characterization of firms that still ground most theories of competitive markets—markets in which all firms are "price takers" that adjust the quantity of goods that they bring to market based on the prevailing market price of their goods or services.

Maximizing Net Income By Selling Goods and Services

In today's commercial societies, essentially everyone participates in markets, rather than a minority of the population. The majority of persons in most places, until a century and a half ago, were farmers or farm workers, or hunters and gatherers. Such specializations were somewhat less dependent on markets for their subsistence than urban workers or village specialists.

Sellers in commercial societies have an interest in maximizing their net income from their commercial activities. This may be tempered a bit by other interests such as a reputation for quality and fairness in dealings with customers, or the value of leisure. Holding those other considerations constant, profit-maximization is a reasonable first characterization of a proprietor's interest in producing goods for sale and an input provider's sale of his or her services.

Since net income is simply the revenues received from sales less the cost of bringing those products to market (including opportunity cost), it is a form of net benefit. Thus, the logic and geometry of the net-benefit maximizing models developed above for consumers also applies to a firm owner's decisions about what to produce and how much to bring to market (e.g. how much to supply).

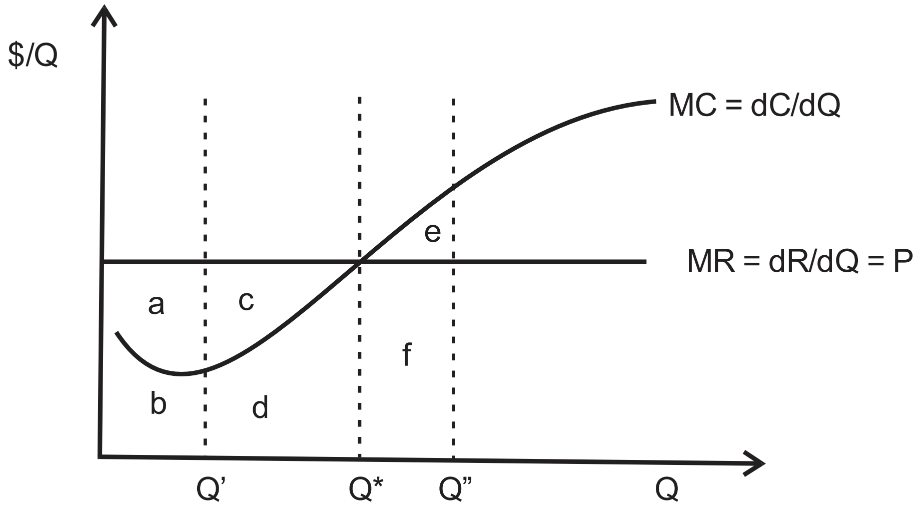
However, in contrast to the consumer's choice, for suppliers, revenue is the benefit, rather than the cost. A firm's cost is determined by the process (technology) used to manufacture or otherwise acquire the goods and services to be sold.

Figure 3.1 illustrates a proprietor's output decision. It assumes a market setting in which firms are "price takers." Such firms can sell all that they want at a given market price. In such a choice setting, the marginal benefit is the same as a firm's marginal revenue, which in turn equals the prevailing market price. The marginal cost of bringing a good or service to market is the firm's expenditure on inputs and the opportunity cost of proprietor efforts used to produce successive units of the product sold.

Notice that the diagram is very similar to the diagram of net-benefit maximizing choice in Chapter 2, except that a firm's marginal benefit curve (MB or MR curve) is now a horizontal line equal to price, and its marginal cost curve (MC) is an upward sloping line reflecting diminishing marginal returns in production.

As before, the net benefits are maximized where marginal benefit (here marginal revenue) equals marginal cost. The mathematical requirements for maximizing net benefits remain the same, although the nature of marginal benefits and marginal cost have changed significantly. What differs is the effect of market prices on decisions. Price is the marginal benefit for price-taking firm owners, although it is the marginal cost for price-taking consumers. They are opposite sides of the same transactions. An increase in prices tends to increase net benefits (profits) for firms, but tends to reduce them for consumers—other things being equal.

Figure 3.1: Maximizing Net Revenue / Profits



This mathematical property of profit-maximizing choices can be demonstrated by focusing on the areas under the marginal revenue and marginal cost curves. Recall that the areas under the marginal revenue and marginal cost curves can be used to determine the net benefits (profits) realized at various quantities that may be chosen. We'll focus on an arbitrarily chosen quantity above and below Q*, the quantity at which MC=MR.

At $Q' < Q^*$, net benefits are $(\mathrm{a} + \mathrm{b}) - \mathrm{b} = \mathrm{a}$. At Q* net benefits are $(\mathrm{a} + \mathrm{b} + \mathrm{c} + \mathrm{d}) - (\mathrm{b} + \mathrm{d}) = (\mathrm{a} + \mathrm{c})$, and at $Q'' > Q^*$, net benefits are $(\mathrm{a} + \mathrm{b} + \mathrm{c} + \mathrm{d} + \mathrm{f}) - (\mathrm{b} + \mathrm{d} + \mathrm{e} + \mathrm{f}) = (\mathrm{a} + \mathrm{c} - \mathrm{e})$. Note that the largest net revenue $(\mathrm{a} + \mathrm{c})$ occurs at Q*, where $MR = MC$.

Outputs below Q* have profits that are below that associated with output Q* (areas analogous to c less), and output levels greater than Q* have additional costs that exceed the revenues generated (analogous to area e).

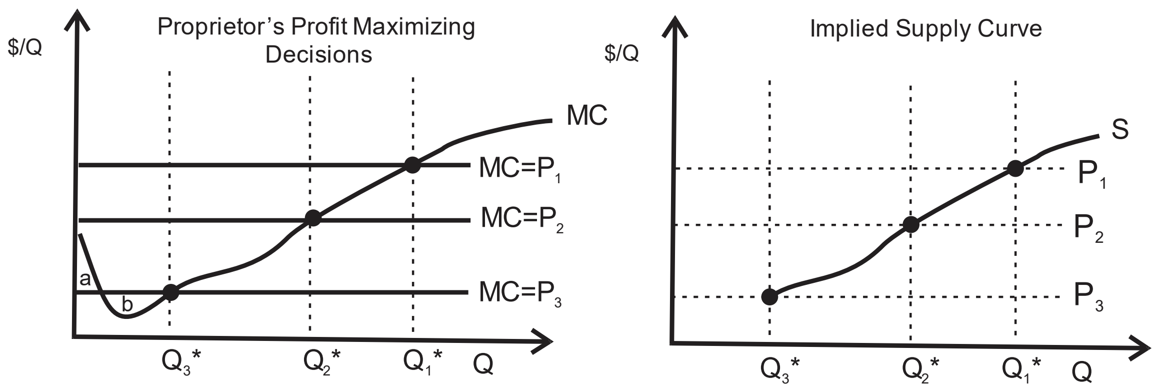
As before, there are also cases where the net-benefit maximizing output is not a positive output, but zero. There are prices that produce marginal revenues that are always below marginal production costs. Thus, there are many goods that are never produced because an insufficient number of consumers are willing to pay a price equal to or greater than the average cost of producing such goods and services.

There are also diagrams in which it may appear that a proprietor or firm would want to produce infinite amounts of the good because the marginal revenue is always above the marginal cost of production—some computer software products resemble such cases—but ultimately wealth, income, and borrowing limits imply the consumers have a limited ability to pay for a good or service. Thus, such infinite supply cases are actually impossible unless there is a complete absence of scarcity and market prices approach or equal zero.

The geometry of deriving a firm's supply curve is also similar to that of a consumer's demand curve. When firms are price takers (e.g. adapt to a prevailing price) they will produce at their net revenue maximizing output levels. These vary with the prevailing market price.

Deriving a firm's supply curve in such cases requires one to vary price and plot the firm's profit-maximizing output level for each price. In most derivations, it's assumed that each firm produces a single good or, if it produces multiple goods, each production line is entirely independent of the others (which is not always true of course). Other prices—including input prices—are assumed to be constant. Figure 3.2 illustrates this process for the firm above.

Figure 3.2: The Geometry of Supply



Outputs for three prevailing market prices are illustrated. The proprietor's or firm's decision in each case is modelled as an effort to maximize net income or profit. In principle one does this for every conceivable price, but in practice the outputs for a few prices are determined and plotted, and the rest are "smoothly" interpolated.

In the case depicted, marginal cost is assumed to fall at first, as the firm tools up for producing the good or service for sale, but after some point (here, fairly early on) marginal costs begin to rise because of diminishing returns in producing the good or service of interest. This "non-monotonicity" assumption has implications for the profit maximizing decision of the firm because there are some outputs where MR=MC, but which yield a negative profit because costs are greater than revenues. Such outputs are never produced (except by mistake).

For this firm, output Q_3* is the "break-even" or "shut-down" output for price P_3. It is the smallest output that this firm will ever rationally produce, and P_3 is the lowest price at which the firm will keep its doors open. At any price lower than P_3, the firm will earn a negative profit on every possible positive output, and so is not able to sustain itself. At the (Q_3*, P_3) price and output combination, the area of loss (area a) exactly equals the area of profit (area b). (Recall that the area under the MR curve from zero to quantity Q_3 is total revenue, and the area under the MC curve is total cost of production or total variable cost, if fixed costs exist.)

The supply curve consists of profit-maximizing outputs for all prices at which profits are greater than zero (or more precisely, greater than the proprietor's opportunity cost rate of return). The supply curve includes the subset of the points on the firm's marginal cost curve at which profits greater than zero are realized (or more precisely, where profits are expected to be realized). Had the marginal cost curve been monotonically increasing, the firm's supply curve would have included all the points on the firm's marginal cost curve for $Q \geq 0$. In such cases, the supply curve goes through all of the same points as the firm's marginal cost curve, rather than a subset of them.

Although a firm's supply curve may include all of the points on its marginal cost curve, they are not the same functions. Marginal cost maps from outputs (Q) into dollars per unit costs (P). Supply maps from prices (dollars per unit) into output levels. The marginal cost function of a firm and its supply function are inverse functions.

The neoclassical theory of the firm often makes a distinction between long run and short run supply. In the short run, there are fixed costs that cannot be avoided or changed. Such costs affect profit levels but not output levels in the short run, because they do not affect marginal (short run) production costs. In the long run, all factors of production can be varied and so there are no fixed costs.

This difference implies that the short and long run marginal cost curves differ, with a firm's long run marginal cost curve being "flatter" than the short-run marginal cost curve, because the proprietor or firm's management has more flexibility in choosing production methods in the long run than in the short run and so can do more to keep marginal costs down. However, the essential geometry of a price-taking firm's output decisions is not affected. In both the short and long runs, it chooses an output that sets marginal revenue equal to marginal cost, or equal to zero if profits greater than or equal to zero cannot be realized.

Although the process through which one derives a firm's short and long run supply curves is identical, the results differ because a different marginal cost curve is used for each. Short run supply is developed from a short run marginal cost curve and long run supply is developed from a firm's long-run marginal cost curve. As a result, a firm's long run supply curves tends to be "flatter" and more price elastic than their associated short run supply curves.

If fixed costs are thought of as the larger more specialized capital employed in production, then there will be only a single point in common between the long run and short runs supply curves, namely, that where the capital stock is perfectly matched to the short-run output of the firm. At outputs greater than that point, the SR-MC curve and its associated supply curve will be above its long run supply curves (and its LR MC curve), because it has less than optimal amounts of capital. At outputs below that level, marginal costs will be lower than their long-run counterpart because the firm has more capital than is optimal for that output.

Characterizing exceptions to the "rule of thumb" that profit maximizing firms "all" produce outputs where MR=MC is generally easier with geometry than with calculus. Most calculus-based models assume that the MC is monotone increasing in output levels, even though it is clear that few if any firms ever produce infinitesimal levels of output. The u-shaped one used in figure 3.2, with its zero-production level well above zero is, thus, more realistic than those normally employed in calculus models.

Nonetheless, the conditions at the margin that determine positive output levels are the same in most such models. Firms maximize profits, they chose output Q* where MC=MR.[1]

[1] Mathematically characterizing possible "corner solutions" is normally done using the Kuhn-Tucker approach, which is beyond the scope of this textbook. There is no appendix on this method because it is cumbersome to use for comparative statics, and so rarely appears in published economic work. Consult a mathematical economics text for more on this method.

III. Deriving Supply Curves from Cost Functions Using Calculus

In Chapter 2, consumers are identified by their marginal benefit or marginal utility functions. Tastes are assumed to be stable during the period of analysis. Similarly, firms are identified by their total or marginal cost functions, which are also assumed to be stable during the period of analysis.

Several other variables are also being held constant when one derives a consumer's demand function, including the prices of other goods and personal income or wealth. Similarly, variables that affect a firm's total cost function are assumed to be constant in order to characterize a firm's supply curves and supply functions. A total cost function includes all the variables that affect costs, including such things as output levels, input prices, technology of production, regulations, taxes on inputs or fines associated with violating safety or environmental laws (either by accident or not), and so forth.

All of the determinants of cost except output levels are held constant, not because they are normally completely stable, but because the aim is to characterize the relationship between the prevailing market price of the product of interest and the firm's output decisions. The other factors also affect output decisions through effects that can be modelled as affecting the shape (as with technology) of a firm's supply curve, or as shifts in its supply curve.

This assumption that other factors are constant (ceteris paribus) does not imply that the other factors are unimportant or are normally constant. Rather they are held constant because a firm's supply function simply maps from prices into outputs. Factors that influence output decisions other than price and output are taken up in the Chapter 4, which demonstrates why supply functions include the prices of inputs and technology as arguments.

The neoclassical theory of price determination (covered in Chapter 5) demonstrates that it is the relationship between market demand and supply functions that ultimately determines market prices. Nonetheless, the other variables affect market prices through their effects on the intersection between market supply and market demand functions in cases in which both firms and consumers are price takers. As demonstrated in Chapter 5, a good deal of mathematics lies behind the geometric supply and demand diagrams taught in principles of economics courses.

Deriving Supply Functions from Exponential Cost Functions

We'll first consider an exponential functional form for a firm's total cost function:

$$C = a + bQ^c \tag{3.1}$$

with C being total cost, Q being the proprietor or firm's output level, and a, b, and c being parameters of the exponential cost function. Assume that a and b are greater than zero and that c is greater than one. The assumptions about a, b, and c assure that costs are greater than zero and that marginal cost rises with output levels reflecting diminishing marginal returns in production. They will also assure that the profit function is strictly concave.

Constant "a" is the firm's fixed costs, which exist only in short run models of cost. Coefficient b can be thought of as roughly the cost of the inputs required to increase output and exponent c as an indicator for the rate at which the required inputs increase as output increases. The latter will vary with the technology of production. Each term in this simple cost equation, thus, has an implicit economic meaning or implication. Together, b, c, and Q jointly determine marginal cost as developed below.

Revenue for a proprietor or firm selling goods in a market in which he or she is a "price taker" is simply $R = PQ$, where R is revenue, P is the prevailing market price, and Q is his or her output level. His or her net revenue or profit, $\Pi,$ is simply total revenue less total cost:

$$\Pi = PQ - (a + bQ^c) \tag{3.2}$$

The output that maximizes profit, Q*, can be found by differentiating the profit function with respect to quantity and setting the result equal to zero.

$$\frac{d\Pi}{dQ} = P - bcQ^{c-1} = 0 \tag{3.3}$$

The first term (P) is marginal revenue. Total revenue increases by amount P every time another unit is produced and sold. The second term (bcQ^{c-1}) is marginal cost. It is the rate at which total cost increases as output increases.

Equation 3.3 implies that a firm owner that is interested in maximizing profits should choose an output where marginal revenue (P) equals marginal cost (bcQ^{c-1}), as in the geometric case illustrated in figure 3.1. Note that fixed costs (a) plays no role in short run output decisions whenever net revenues at Q* are greater than zero (ignoring fixed costs). If fixed costs (a) are greater than short run net revenues, the firm may choose to go out of business, but otherwise fixed costs do not affect output choices in the short run. (Closing the firm is normally regarded to be a long-run choice.)

To be sure that this "first order" condition characterizes a maximum rather than a minimum, the second derivative of 3.2 should also be calculated and its "sign" determined (e.g. whether it is always greater than zero, less than zero, or may have different "signs" as Q varies). This is possible for most concrete function forms of profit functions. In this case, the second derivative is:

$$\frac{d^2\Pi}{dQ^2} = -bc(c-1)Q^{c-2} < 0 \qquad (3.4)$$

Notice that b, c, (c-1), and Q* are all greater than zero, so the slope of the marginal cost curve ($bc(c-1)Q^{c-2}$) is positive as is consistent with diminishing marginal returns. The minus sign in front of that slope implies that the entire term is negative for any quantity greater than zero. Thus, the profit function is strictly concave in the domain of interest.

Note that the concavity of the profit function is entirely determined by the shape of the cost function in settings in which firm owners sell their product in markets at "given" prices. (Marginal revenue in the "price taker" case is flat (MR=P) and thus, the marginal revenue function has zero slope and does not affect the magnitude of the second derivative of the profit function.)

To characterize the supply curve, we solve equation 3.3 for Q as a function of market price. A few algebraic steps can do so. First shift all the terms with Q in them (here just one) to the lefthand side of the equation by adding bcQ^{c-1} to each side.

$$bcQ^{c-1} = P \qquad (3.5)$$

Divide each side by *bc*, and then raise each side to the $\frac{1}{c-1}$ power. These steps generate:

$$Q^* = (\frac{P}{bc})^{1/c-1} \qquad (3.6)$$

This is the **supply function** for the proprietor's or firm's output decision being modelled.

Note that the amount produced and sold increases with price. (Price is in the numerator.) Supply diminishes as "c" increases, which is to say, as the rate at which additional inputs produce additional outputs increases. It also falls as the cost of inputs (b) increases. Thus, every supply function from this family of exponential cost functions has the property that economic intuition suggests a supply function should have—and conversely, the results also show that economic intuition about a firm's supply decisions is consistent with at least one fairly general family of cost functions.

This illustration demonstrates that basic economic intuitions about supply are not inherently "wrong" or logically impossible—although it is possible that they are reliable only for a subset of possible cost functions.

Deriving Supply Functions from More General Cost Functions

More abstract and encompassing families of cost functions can also be used to characterize a firm's supply function. For example, assume that the firm's cost function is simply:

$$C = c(Q, w, r) \quad (3.7)$$

with C being total cost, Q being the firm's output level, *w* being the prevailing wage rate for labor and *r* being the cost of capital. Assume that the first derivatives of the cost function are all positive and the second derivatives are also all positive (to reflect diminishing marginal productivity). The cross partials are also assumed to be positive. Total cost naturally rises as output increases and if input prices increase. The assumptions about second derivatives will help to assure that cost function is convex and that the profit function is strictly concave. The concavity of the profit function in the Q dimension is most important for the derivation of the slope of a firm's supply function, as demonstrated below. Chapter 4 will demonstrate why input prices belong in a firm's cost function—although it should be plausible for most undergraduate economics majors.

Revenue for a proprietor or firm selling goods in a market in which he or she is a "price taker" is simply $R = PQ$, where R is revenue, P is the prevailing market price, and Q is his or her output level. The proprietor's or firm's profit, $\Pi,$ is simply total revenue less total cost:

$$\Pi = PQ - c(Q, w, r) \quad (3.8)$$

The output that maximizes profit, Q*, can be found by differentiating the profit function with respect to quantity and setting the result equal to zero.

$$\frac{d\Pi}{dQ} = P - dC/dQ = 0 \quad \text{at } Q^* \quad (3.9)$$

The first term (P) is marginal revenue. Total revenue increases by amount P every time another unit is produced and sold. The second term (dC/dQ) is marginal cost. It is the rate at which total cost increases as output increases. Thus, equation 3.9 implies that a firm owner that is interested in maximizing profits should choose the output where marginal revenue (P) equals marginal cost ((dC/dQ), as in the geometric case examined earlier in the chapter.

To be sure that this "first order" condition characterizes a maximum rather than a minimum, the second derivative of equation 3.8 should also be calculated, and its "sign" determined. This is possible for most concrete function forms of profit functions. In this case, the second derivative is:

$$\frac{d^2\Pi}{dQ^2} = -\frac{d^2C}{dQ^2} < 0 \tag{3.10}$$

We have already assumed that $\frac{d^2C}{dQ^2} > 0$ in order to characterize the effect of diminishing marginal returns in production and assure that the profit function is strictly concave.[2] The profit function is strictly concave in the ΠxQ plane, because the marginal cost function is upward sloping. As in the previous model, marginal revenue in the "price taker" case is a horizontal line (MR=P), and thus, the marginal revenue function has zero slope and does not affect the magnitude of the second derivative.

Keep in mind that the marginal cost function, $\frac{dC}{dQ}$, implicitly includes the same argument as its "parent" function, which is to say, it is a function of Q, w, and r, just as the total cost function was. The conventional notation hides this, but whenever dealing with abstract functions, it has to be kept in mind, because it affects a variety of calculations that can be undertaken with such functions, including the step taken next to characterize the firm's supply function and its slope.

To characterize the supply curve, we make use of the implicit function theorem. Recall that implicit function theorem implies that any (locally) differentiable function that has the value zero at all relevant points has the property that each variable can be described as a function of every other. Note that equation 3.9 characterizes a function this type. It can be written as H= h(P, Q*, w, r) = 0. Thus, we use that equation to characterize Q* as

---

[2] Sometimes this second derivative will be written as simply $\frac{d^2\Pi}{dQ^2}$. This is either being used as a shorthand for the longer expression, or implicitly assuming that the cross partials are all zero. Notice if the wage rate and interest rates terms were left out of the cost function, $\frac{dC^2}{dQ^2}$.would be all that is necessary.

$$Q^* = s(P, w, r) \qquad (3.11)$$

Equation 3.11 is the **firm's supply curve** and the letter given to the function characterized was chosen with this in mind. "s" is intended to remind the reader that this is a supply function. The amount produced and sold varies with price and with the price of inputs.

Next, recall that the implicit function differentiation rule states that if H≡ h(P, Q*, w, r) = $P - \frac{dC}{dQ} = 0$, then $\frac{dQ^*}{dP} = \frac{\frac{dH}{dP}}{-\frac{dH}{dQ^*}}$. We can use that rule to characterize the slope of the supply curve for the firm being modelled. In this case, a relatively simple expression characterizes the slope of the firm's supply function:

$$\frac{dQ^*}{dP} = \frac{1}{-\left(-\frac{dC^2}{dQ^2}\right)} > 0 \qquad (3.12)$$

The derivative of equation 3.9 with respect to P which generates the numerator is just 1. The derivative of equation 3.9 with respect to Q generates the denominator. It is simply the second derivative of the marginal cost function, which has already been reported as equation 3.10. The profit function is concave, which implies that the term inside the parentheses is negative. The minus sign preceding the term inside the parentheses implies that the denominator is positive, as is the numerator.

Thus, the firm's supply curve is upward sloping. As sales prices increases, the firm will produce and sell more units of the good of interest. *This is true of every firm that successfully maximizes profits and has a cost function that exhibits increasing marginal costs at the output chosen.*

Again, we find that economic intuitions based on geometric models of supply are not "wrong" or logically impossible—but rather are very general—indeed more general than might have been expected, because they apply to a very broad family of possible cost functions.

IV. Short-Run and Long-Run Market Supply

To derive a short-run market supply function one adds up the supply curves of every firm that is presently in the market. In the short run there are a finite number of firms in the industry, M. Thus given, supply curves similar to those of 3.11, short-run market supply can be characterized as:

$$S^{SR} = \sum_{J=1}^{M} s_j(P, w, r) \qquad (3.13)$$

Insofar as each firm's supply curve is upward sloping, the sum is also upward sloping.

To derive a long-run market supply function is more complicated, because one cannot always assume that the number of firms currently present in the market will continue into the future. For example, if rates of return (profits) are higher than normal, other firms may enter until the rate of return falls to the usual (opportunity cost) rate of return. Indeed, in the Marshallian approach to long run supply, this is the main adjustment mechanism.

During the late nineteenth century and early twentieth century, which was the period in which neoclassical economics was worked out—the majority of markets were served by relatively small family-based enterprises. The developing West consisted of nations of shopkeepers, with only occasional large-scale enterprises. These small firms were mainly organized as proprietorships and partnerships. Such businesses had limited ability to grow because of the manner in which they were organized and financed. There were limits in the number of employees that they could organize, trust, supervise, and monitor. There were also limits in their ability to raise capital.

Thus, efficient firms were of limited size and Marshall believed that most long run supply adjustments in such markets took place through the entry and exit of efficient-sized firms. Firms "exit" when they stop producing the product of interest and start producing others or become bankrupt. Firms would enter when they left some other market or a proprietor or partnership organized a new business.

Because shopkeepers used similar methods to produce and sell their products and services, they tended to be of roughly the same size. As a first approximation, one could think of them as identical in size.

In such cases, long-run market supply at a given price and level of demand was simply "N"—the number of efficient-sized firms necessary to meet demand at that price. In the long run, Marshall argued, and most other neoclassical models of perfect competition agreed, entry and exit would continue until profits equaled the typical proprietor's opportunity cost rate of return. This is sometimes referred to as a zero profit equilibrium, when opportunity cost rather than production cost is used to determine net returns from a given enterprise. And long-run supply functions tended to be horizontal lines equal to the minimum average cost of production by efficient sized firms. In the long run, firm owners realized only the normal rate of profit on their investments, time, and energy.

In today's commercial societies, non-chain coffee shops, bicycle shops, restaurants, and clothing stores resemble the Marshallian market environment. They are all roughly the same size. Entry and exit constantly take place, and the average firm earns approximately the "normal" rate of return (adjusted for risk) in the market of interest. When demand increases, such firms may expand a bit, and many other small-scale entrepreneurs enter to meet the new demand. When times are tough, because demand is relatively low or input prices are relatively high, many firms exit.

However, entry is not always as easy as assumed in the Marshallian model. Entry barriers of several kinds may exist as consequence of scarce factors such as land close to a city center, entrepreneurial ability, zoning, patents, or licensing requirements that limit or slow entry. Moreover the "best" locations may already be taken by existing firms. In addition, concentrated ownership is encouraged by advantages associated with diversification across locales that make regional chains of cafes, bicycle shops, hair salons, restaurants safer investments than that available to businessmen and women with only single storefronts.

An alternative to the Marshallian model, which this author prefers for many purposes, is one grounded in Ricardo's (1817) observations about farming in the region around London. Farms with especially fertile land close to a city center earn higher returns (command higher rents) than those farther away or with less fertile land. Differences in fertility and location affected the cost of raising crops and taking them to market. Thus, rental rates were higher for some parcels than available to otherwise similar properties with less fertile or more distant locations.

Both factors limited entry into the agricultural market around London. In such cases, market prices determine the marginal farmer—the one whose distance from market and land fertility are just sufficient to support production, while other landholders closer to London or with more fertile land realize rents (higher than average profits) from their land.

A similar logic holds for businesses where location, differential access to inputs, and patent or copyright protection imply that production costs vary among producers. In such cases, the marginal producer earns an ordinary return on their investments, while other long-term owners of land realize above normal profits.[3]

---

[3] It is sometimes argued that the proper way to determine excess profits is the value of a property on the market. Insofar as this tends to be the present discounted value (see Chapter 6) of the profits associated with a property or business, such calculations imply that firm owners earn "only" their opportunity cost rate of return, which is risk adjusted interest rate used for such calculations. By focusing on realized net income, the analysis uses profit in its more common meaning

In a Ricardian market, every firm uses a somewhat different production function and therefore has a cost function that differs from other firms serving the same market. As a consequence, firms have different short and long-run supply functions.

In Ricardian markets, entry is limited and long-run supply is the sum of each potential supplier's long run supply function. There is no presumption that every firm is of roughly the same size, nor that entry takes place until all firms realize average profits (their normal rate of return on investment, adjusted for risk). As in agriculture, there may be a sufficient number of suppliers that each acts as a price taker; which is to say that each firm acts as if its supply decisions will have no effect on market prices, as assumed in the models of supply developed to this point in the chapter.

In such cases, the logic and derivation of long-run supply is very similar to that for short run supply. If there are M firms that potentially could produce the product or service of interest, the long-run supply function for that market is the sum of the individual long run supply functions.

$$S^{\mathrm{LR}} = \sum_{J=1}^{M} s_j^{LR}(P, w, r) \qquad (3.14)$$

Obvious examples of such markets include many natural resources. Iron, copper, gold and oil deposits differ in size, quality, and location in a manner that affects production and transportation costs. And as Ricardo noted, a mine or oil well may be closer to market, be more productive or be in a region with more fortuitous weather. Similarly, a firm may have favorable access to inputs, an especially talented workforce, or an entrepreneur who is able to keep costs down and quality up.

Similar effects may be associated with relatively small franchise organizations, as with local chains of restaurants, that compete with each other, but have very similar menus and service. Profits will vary among firms (and franchise outlets), but only the last or smallest firm will earn zero profits (e.g. have net revenues equal to their firm owners' opportunity cost rate of return).

For Ricardian markets the area-based diagrams of the first part of the chapter can be applied for both long run and short run analysis. For the Marshallian case, long and short run

in ordinary usage. The former implies that it is impossible to realize excess returns in well-functioning, well-informed complete market systems. The latter allows such possibilities to be recognized. That even in well functioning markets, individuals may disagree about the value of, for example, stocks is evidenced by the trades that take place every day and from systematic differences in rates of return among investors.

diagrams differ, with long run supply always being a horizontal line with each firm producing at the minimum of its long-run average cost curve (a point similar to Q_3 in figure 3.2).

Although competition may be intense in Ricardian markets in the same sense that it is in Marshallian markets, profits may vary widely among firms in such markets. Only the marginal firms earn their opportunity cost rate of return. The others earn different levels of profit depending on their costs. Moreover, their owners may strictly prefer their particular firm and industry to others available to them, rather than being indifferent among them as in the Marshallian analysis.

V. The Geometry of Supply by Firms with Price-Setting Abilities (Monopoly Pricing)

Firms are not always price takers. In some places and in some markets, there may be only a single firm or a small number of firms that produce a product or service. The purchasing opportunities provided by the internet has reduced the number of products where this is true, but many products and services still exist where there is only a single supplier for a good or service. For example, in moderately sized towns, there may be only a single coffee or bicycle shop, or a single restaurant that provides meals based on particular regional cuisines. There are many towns where there is only a single restaurant that sells meals from the French, Indian, Chinese, or Mexican cuisines.

All firms compete with other firms for the purchases of local consumers, but firms that sell unique products are price makers rather than price takers. They face their own market demand function, one whose shape is affected by available substitutes and their prices, but which is nonetheless unique. Their supply and pricing decisions thus affect the level of demand for their products.

They do not directly alter their market demand function, which is determined by consumer preferences, relative prices, and income; instead, they influence the quantity purchased by their consumers by setting the price their consumers pay for their products and services. Such effects imply that production costs and demand jointly determine their profits (net revenues). In this case, price is one of the variables controlled by the firm, rather than an exogenous variable over which they have only a trivial influence.

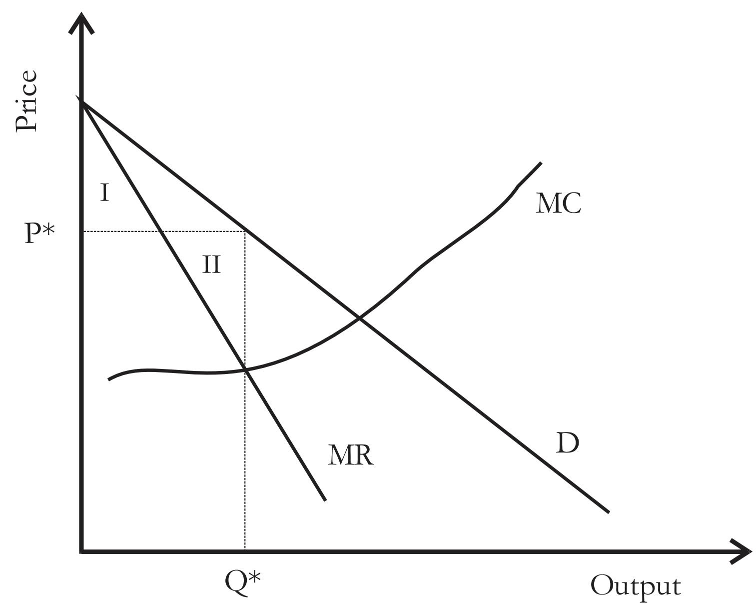
Similar price setting ability also occurs in a subset of broader markets, where economies of scale, high rates of innovation, and patent or copyright protection may cause some types of products to be produced by only a single firm or small group of producers.

Price-making merchants often each have their own downward-sloping demand curves for their products. In principle, such firms may charge any price that they want, but their sales are ultimately determined by their consumers, rather than the firm.

Neoclassical models of monopoly pricing imply that the output decisions of such firms take account of how their pricing decisions affect sales as well as the marginal production costs of goods or services sold. They do not charge the highest price at which their product can be sold, but rather the price and output combination that maximizes profits.

Figure 3.3 illustrates the geometry of the output decision of such a firm. The fact that a firm faces a downward sloping demand curve—no matter how flat or sharply downward sloping—implies that its marginal revenues are no longer determined by a horizontal line equal to price. (The calculus illustration in the next sub section will make it clearer why that is the case.)

Figure 3.3 Monopoly Output and Pricing



Given the results from chapter 2 and the first half of the present chapter, we know that this firm will choose an output where the marginal cost of production equals the marginal revenue generated by it. The product is priced so that all that is produced sells, so the price in this case is that at which demand will exactly equal supply, which is the one labeled P*. However, marginal revenue is no longer equal to price.

The demand curve in Figure 3.3. is assumed to be "linear" (e. g. a straight line). In that case, it turns out that the marginal revenue is also a straight line. In particular, marginal revenue is a line that lies halfway between the demand curve and the vertical axis. (The calculus in the next subsection will make that relationship clear.)

The area that corresponds to the firm's profits can be calculated in two ways. First, one can use the "area" approach developed in Chapter 2 and applied to firms in the first half of this chapter. The firm's total revenue at Q* is the area under the marginal revenue (MR) curve from 0 to Q*. Its total (variable) cost is the area under its marginal cost curve from 0 to Q*, and its profits are the difference between them. In the diagram it is the somewhat irregular shaped triangle between the marginal cost (MC) curve and the marginal revenue curve between 0 and Q*.

Another area that corresponds to total revenue is simply the rectangle whose area is P*Q*, which turns out to be the rectangle characterized by the point (P*, Q*) on the demand curve and the lines that go directly to the horizontal and vertical axes. Using this measure for total revenue generates another area for profits which is the somewhat irregular area between the marginal cost curve and the horizontal line equal to P*. Since the two profit measures have to be the same (the firm only earns one profit, a real number) it turns out that the triangles labeled I and II have exactly the same areas. (In fact, they are identical congruent triangles in the case depicted.)

This is just a geometric regularity of interest only because of the different ways that total revenue can be calculated in Figure 3.3. That feature of the diagram does not have much economic meaning, although it is a useful way to determine whether you really understand the geometry of this diagram or not.

VI. Output Choices by "Price Making"

To illustrate the mathematical model and calculus that lies behind Figure 3.3, consider the case where a firm faces a downward sloping demand curve Q = a – bP and has a total cost function $C = zQ^t$ where z>0 and exponent "t" is greater than or equal to 1.

To characterize a firm's total revenues in terms of output, we'll need to first find the inverse of the demand function, which is to say the function describes the price P at which output Q can be sold given the demand curve. This requires solving $Q = a - bP$ for price as a function of output. Subtracting "a" from both sides of the equation and dividing by "-b" generates the equation $P = (Q-a)/(-b) = (a-Q)/b$, which describes the price at which output Q can be sold. This is the inverse of the demand function.

Total revenue is just PQ as before. However, in this case P is affected by the firm's output decision. Using the equation that we just derived which describes the price at which Q units can be sold, allows total revenue to be written as a function of the firm's output:

$$PQ = [(a-Q)/b]Q = (aQ-Q^2)/b.$$

Next, we write down the profit function that takes account of the effect of the firm's output on both revenues and costs.

$$\Pi = TR - TC = (aQ - Q^2)/b - zQ^t \quad (3.16)$$

To find the profit maximizing output, differentiate the profit function with respect to Q and set the result equal to zero, which characterizes the "first-order condition" for the profit-maximizing output.

$$(a-2Q)/\mathrm{b} - \mathrm{t}zQ^{t-1} = 0 \quad (3.17)$$

Notice that the first term is marginal revenue. It characterizes the rate at which total revenues increase as output increases—including the effect of greater output on selling price. Note that it falls at twice the rate that the demand curve did, at 2Q/b rather than Q/b, which accounts for the fact that the MR curve in figure 3.3 was halfway between the demand curve and the vertical axis. The second term is marginal cost. It characterizes how total costs increase as output increases. The profit-maximizing output occurs at the output, Q*, where MR = MC.

To make sure that the first order condition characterizes a maximum rather than a minimum or inflection point, we take the second derivative of the profit function. If it is always less than zero (negative), then the profit function is strictly concave, and if Q* exists, it will be the profit maximizing output. The second-order condition for a maximum is the derivative of equation 3.17 which is:

$$-2b - (t-1)tzQ^{t-2} < 0 \quad (3.18)$$

Which is less than zero for all quantities greater than zero. "b" is positive, so the first term is negative. (t-1), t, and z are all positive numbers for t>1, so the second term is also negative (because of the leading negative sign). When t=1 (the constant marginal cost case used in equation 3.19), the second term equals zero, but the condition still holds via the first term ($-2/b < 0$). In either case, the profit function is strictly concave. Thus, we can use equation 3.17 to characterize the firm's profit-maximizing output.

With concrete functional forms, one can often—but not always—solve the first-order condition (here, equation 3.17) for an equation that characterizes Q*, which is the output or **supply function for a price-making firm**. In this case, however, such a solution is not possible for the general class of exponential cost functions used to this point.

Nonetheless, a few special cases from the assumed family of cost functions can be solved. To illustrate such a case, suppose that exponent t=1; in other words suppose that the firm's cost function has constant marginal costs. In this case equation 3.17 takes the form:

$$(a - 2Q^*)/b - z = 0 \qquad (3.19)$$

And, a bit of algebra allows Q* to be characterized as a function of parameters of the firm's choice setting. First, add 2Q*/b to both sides to obtain:

$$a/b - z \ = 2Q^*/b$$

Next, reverse the sides of the equality, and multiply both sides by b/2 to get:

$$Q* = \ (b/2)\,[a/b - z] \ = a/2 - bz/2 = \ (a - bz)/2 \qquad (3.20)$$

Equation 3.20 is the **price-making firm's supply function**. It characterizes the firm's output in terms of parameters of the demand function (here a and b) and its cost function, (here z). Note that it does not map from prices into outputs, because price is "endogenous" in this case. Rather it summarizes how this firm's output varies with parameters of its demand and cost functions.

By inspection, we can see that the higher "a" is the greater is output. However, the greater are b and z, the lower is output, other things being equal. If we knew precise values for a, b, and z, we would know the firm's precise output.

Its selling price can be determined by substituting Q* into the inverse demand equation worked out at the beginning of this section. In the case modeled, the equilibrium price is:

$$P^* = (a-Q^*)/b = (\ a - [\tfrac{a-bz}{2}])/b \qquad (3.21)$$

An Abstract Model of Output Choices by a Price Making Firm

A more general model of a "price-making" firm's decision can also be developed using more general characterizations of a firm's demand and cost functions. We can assume that a firm faces the demand curve, $Q = q(P, Y, N, P_j)$ where P is the price of the firm's product, Y is average consumer income, N is the number of consumers in this market, and Pj is the price of a substitute or complement for the firm's product. Its cost function can be characterized as $C = c(Q, w, r)$ where Q is the firm's output and w is the wage rate paid for its labor and r is its rental cost for capital equipment. Demand functions similar to this one were worked out in Chapter 2, cost functions similar to this one were used earlier in this chapter, the basis for which are worked out in the next chapter. Profits are again simply total revenue less total costs, or $\Pi = PQ - C$.

To characterize the profits as a function of output, we'll again need to find the inverse demand curve. Notice that a "zero function" can be created by subtracting the demand function from the desired output by consumers, $Q - q\left(P, Y, N, P_j\right) = 0$. That allows us to use the implicit function theorem to describe sales price as a function of output, $P = p(Q, Y, N, P_j)$.

Substituting that relationship into profit function yields:

$$\Pi = p(Q, Y, N, P_j)Q - c(Q, w, r) \qquad (3.22)$$

To simplify a bit, it will simply be assumed that the profit function is strictly concave. Given this, the first order condition will characterize the firm's profit maximizing output. The first order condition is simply:

$$\frac{d\Pi}{dQ} = \left(\frac{dP}{dQ}\right) Q + P - \frac{dC}{dQ} = 0 \equiv H \qquad (3.23)$$

The first two terms, $\left(\frac{dP}{dQ}\right) Q + P$, are the firm's marginal revenue function. (Note that we've used the rule for differentiating products to find the firm's marginal revenue function.) The last term is the firm's marginal cost function, $\frac{dC}{dQ}$. So again, the firm's profit maximizing output, Q*, is where marginal revenue equals marginal cost.

At Q*, the first derivative of the profit function equals zero and so the implicit function theorem can again be applied to characterize any variable of interest in terms of the others in the revenue and cost function being assumed. We are, of course, most interested in Q* and we'll want to write down the function that describes it.

$$Q^* = s(Y, N, P_j, w, r) \tag{3.24}$$

This characterization of the firm's output reveals that a firm's output is jointly determined by parameters of the firm's demand and supply functions. Y, N, and Pj are parameters of its demand function. And w and r are parameters of its cost function.

The implicit function differentiation rule can be used to characterize the effects of changes in those variables on the firm's output. If we define equation 3.23 as H, the partial derivative of Q* with respect to wage rates can be written as:

$$\frac{dQ}{dw} = \frac{\frac{dH}{dw}}{-\frac{dQ^*}{dQ}} = (-\frac{\mathrm{dC}}{\mathrm{dQdw}})/(-\frac{d^2\Pi}{\mathrm{dQ}^2}) < 0 \tag{3.25}$$

The profit function has been assumed to be strictly concave, thus $\frac{d^2\Pi}{\mathrm{dQ}^2} < 0$ and the denominator is a positive number. The term $\frac{\mathrm{dC}}{\mathrm{dQdw}}$ is the effect of an increase in wage rates on marginal cost which will be positive. The higher wage rates are, the higher is the marginal cost of production. The negative sign before that term implies that the numerator is negative, so the price-making firm's output falls as wages increase.

The derivatives with respect to the determinants of demand are a bit more complicated because we have used the implicit function theorem to characterize prices as a function of outputs, thus the implicit function theorem would have to be applied to the zero equation that we used to characterize the price-output relationship. As a consequence, many more terms would be involved.

For the purposes of this chapter, the main point of interest for this relatively general derivation is that the supply decisions of firms with a bit of monopoly power are always jointly determined by the factors that determine their cost functions (as also true of price taking firms) and the factors that determine the extent of its demand. The latter ordinarily includes such considerations as average consumer income, number of consumers, and the prices of substitutes and/or complements to the firm's output.

Although these choices are more complex than those of price-taking firms, they do provide our first explanation of how a subset of market prices for final goods are determined.

The inverse demand curve implies that the **market price** associated with output Q* is simply:

$$P^* = p(Q^*, Y, N, P_j) \tag{3.25}$$

In markets where firms have a degree of price-setting ability, there is no reason to rely on unmodelled minor adjustments in prices by firms or by an invisible Walrasian auctioneer. In such cases, the neoclassical model implies that both output and prices reflect the requirements of profit maximization.

How prices are determined in markets where firms are all price takers is taken up in Chapter 5. Other factors, left out of the core models that may also influence pricing decisions are taken up in parts II and III.

VII. On the Generality and Limits of the Neoclassical Theory of Supply

As true of the consumer choice model developed in Chapter 2, the neoclassical model of a firm's output decisions developed in Chapter 3 is more general and has broader implications than it might at first appear. First, it demonstrates that firm decisions are largely driven by their marginal costs. Both marginal production costs and marginal opportunity costs inform the decisions of firms. These together with market prices and market demand determine which goods get produced and the extent of their respective sales to consumers. In all cases, a profit-maximizing firm will choose the output that sets its marginal costs equal to its marginal revenues.

In cases where firms are price takers, firms respond to prevailing market prices by producing the quantity of their product that maximizes the firm's profits at the prevailing market price. Firms in such markets produce at a sufficiently small scale that their decisions do not affect market prices. In cases in which a firm has price making abilities, it will also produce outputs where marginal cost equals marginal revenue, but marginal revenue no longer equals price and price is not an exogenous variable from the perspective of such firms.

In both cases, marginal costs and marginal revenue considerations determine the output decisions of profit-maximizing commercial organizations. Deviations in outputs from the level that sets marginal cost equal to marginal revenue can only occur by mistake or a lack of information about either the prevailing market price or the firm's production costs.

The analysis also has implications about long-run adjustments in supply in otherwise stable settings. Marshall's model of long run supply implies that in markets with relatively small firms using similar production methods, unusually large returns on investment cannot endure forever, because in the long run entry tends to increase supply (and reduce prices) to the point where only average rates of return are earned. This condition, as modelled by Marshall and many subsequent economists, has many caveats, the most important of which is the assumption that the cost functions of each firm are essentially identical and that efficient sized firms are small relative to the market of interest.

If cost functions differ in the long run, as in Ricardian settings, the "zero profit" equilibrium applies only to marginal firms. But the idea that profits attract entry and losses lead to exit remains an important insight even in such cases. Without profits, firms will not enter and will not be sustained by market transactions.

The latter implies that some goods will not be provided by markets, because their production costs are higher than the price at which they could be sold. And conversely, it will be products that can be profitably sold (e.g. those whose sales generate positive net revenues) that are produced and sold by firms.

In this manner, production costs and the extent of demand ultimately determine the types of goods available to consumers. As will be seen in the next chapter, prices play significant roles on both sides of this commercial calculus.

References

Alchian, A. S. and Allen, W. R. (1974). *University economics, Elements of inquiry* (third edition). Saddle Hill, NJ: Prentice Hall.

Becker G. S. (1965). A Theory of the Allocation of Time, *The Economic Journal* 75: 493-517.

Baumol, W. J. (1982). Contestable markets: An Uprising in the theory of industry structure. *The American Economic Review*, *72*(1), 1-15.

Browning, E. K., & Zupan, M. A. (2020). *Microeconomics: theory and applications*. John Wiley & Sons.

Buchanan, J. M. (1979). *Cost and Choice: An Inquiry into Economic Theory.* Chicago: University of Chicago Press.

Ekelund Jr, R. B., & Hébert, R. F. (2013). *A History of Economic Theory and Method*. Waveland Press.

Encaoua, D., & Jacquemin, A. (1980). Degree of monopoly, indices of concentration and threat of entry. *International Economic Review*, 87-105.

Marshall, A. (1890). *Principles of Economics*. New York: Macmillan and Company.

Muellbauer, J. (1974). Household Production Theory, Quality, and the" Hedonic Technique". *The American Economic Review*, *64*(6), 977-994.

Ricardo, D. (1817) *Principles of Political Economy*. London: John Murray.

Samuelson, P. A. (1947). *Foundations of Economic Analysis.* Cambridge: Harvard University Press.

Shephard, R. W. (1953). *Cost and Production Functions.* Princeton: Princeton University Press.

Stigler, G. J. (1963). Competition and the rate of return. In *Capital and rates of return in manufacturing industries* (pp. 54-71). Princeton University Press.

Stigler, G. J. (1951). The division of labor is limited by the extent of the market. *Journal of political economy*, *59*(3), 185-193.

Varian, H. R. (1992). *Microeconomic Analysis* (3rd ed.). New York: W. W. Norton.

Viner, J. (1931). Cost curves and supply curves. *Zeitschrift für Nationalökonomie*, *3*, 23–46.